

⑬ 日本国特許庁 (JP)
⑫ 公開特許公報 (A)

⑪ 特許出願公開

昭57—86899

① Int. Cl.³
G 10 L 1/00

識別記号

庁内整理番号
7350—5D

③ 公開 昭和57年(1982)5月31日

発明の数 1
審査請求 未請求

(全 4 頁)

④ 音声認識装置

② 特 願 昭55—163143
② 出 願 昭55(1980)11月18日
⑦ 発 明 者 後藤正彦
尼崎市南清水字中野80番地三菱

電機株式会社通信機製作所内
⑦ 出 願 人 三菱電機株式会社
東京都千代田区丸の内2丁目2
番3号
⑦ 代 理 人 弁理士 葛野信一 外1名

明 細 書

1、発明の名称

音声認識装置

2、特許請求の範囲

(1) 入力単語音声と登録単語音声と比較し登録単語音声を識別する認識手段、

前記入力単語音声の音節数を検出する検出手段、

前記登録単語音声の音節数を記憶する記憶手段、

および

前記入力単語音声の音節数と上記登録単語音声を比較し、認識動作において入力単語音声と同一または近似の音節数をもつ登録単語のみを照合対象とする比較手段を備えた音声認識装置。

(2) 前記認識手段は、音声分析・特徴抽出回路と特徴パターン圧縮回路を含む特許請求の範囲第

(1)項記載の音声認識装置。

(3) さらに登録モードにおいて前記検出手段を動作させ認識モードにおいて前記比較手段の動作の結果入力単語音声と登録単語音声が同一または近似の音節数を有するとき前記認識手段を動作さ

せる登録／認識モード切替手段を備えた特許請求の範囲第(1)項記載の音声認識装置。

3、発明の詳細な説明

この発明は音声認識装置に関し、特に認識処理時間の高速化を図った単語音声認識装置に関する。

音声自動認識に関し多くの実験がなされ、近年実用化段階に至る開発が急速度で展開されている。

第1図は従来の単語音声認識装置の一例を示す。マイクロホン1から入力された音声波形2は、音声分析・特徴抽出回路3内で例えば周波数スペクトラム分析を受け、スペクトラムの時間構造を表わす特徴パターン4として抽出される。この特徴パターン4は、次段のパターン圧縮回路5に導入され、ここで発声時間の長短に関係なく一定長の圧縮パターン6に凝縮される。この例では、単語音声の始端と終端間の特徴パターンを等分割し、一定数の標本値に圧縮する線形変換等公知の手法を使用している。

後段のスイッチ7は学習／認識モードを切替えるためのもので、学習登録時には点線の位置、認

BEST AVAILABLE COPY

識実行時には実線の位置に切換わる。認識動作に先立つて使用される単語の圧縮パターン6は、スイッチ7の点線位置を介して登録パターンメモリ9に順次記憶させる。一連の例えば100語分の単語の登録動作が終了すると、スイッチ7は実線位置に切換えられる。以後の発声による圧縮パターン6は入力パターンメモリ8に一時記憶される。入力パターンメモリ8は発声毎に更新し書き換える。認識判定回路12によつて入力パターン10の全体に対して順次複数の登録パターン11と比較し、最も類似度の高いパターンを選択しこれを認識結果13として出力する。

しかし、従来のこの種の音声認識装置においては認識処理時パターン照合動作を登録単語数だけ繰返し実行する必要がある、多くの時間を要していた。例えば、1単語分の照合に5 msecが必要であると仮定すると、100語では0.5 sec、1000語では5 secにも達し、実時間応答性を著しく損うものであった。この欠点は登録単語数の増加に比例して助長される。

音声分析・特徴抽出回路3内で抽出された特徴パターン4は音節計数回路14にも入力される。音節計数回路14は入力単語音声の音節数を計数する。例えば、「レー」なら「1」、「イチ」なら「2」のごとく計数する。なお、音節に限らず、これに類する例えば音素数、モーラ数等を計数するように構成してもよい。公知のごとく、音声字母は概して音波のある時間的部分とも考えられる区分をその音価に着目して分類される。音節(syllable)は単独で発声の単位となり得る最小の音形単位を表わす。一方音素(phoneme)は音韻表記で用いられる一つ一つの字母で表現される音路の仮想的な時間区分的音形単位を表わす。この音素の概念の中には、独立した音節をなさないが、時間区分としては一つの単位をなすような長音節の末尾要素(例えば、「オーオトコ」)が存在する。このような時間区分的単位は一般にモーラ(mora)数と呼ばれる。以下の説明および特許請求の範囲においてこれらの音形単位を音節と言う。

得られた音節数15は登録モードにおいて学習

そこで、この発明は特徴パターンの他に単語音声の音節数またはこれに類する情報(例えば、音素数、モーラ数等)も併せて登録し認識処理時に入力単語音声と同一の音節数またはこれに類する情報をもつ登録単語以外を認識動作の対象から除外することによつて、処理時間の短縮化を図った音声認識装置を提供することを目的とする。

この発明は、要約すれば、入力単語音声の音節数またはこれに類する情報を計数する手段および前記情報を蓄積比較する手段を設け、認識処理時に入力単語音声と同一または近似の音節数またはこれに類する情報をもつ登録単語のみを照合対象に限定し認識処理時間の高速化を図ったことを特徴とする。

以下に、図面に示す実施例とともにこの発明をより詳細に説明する。

第2図はこの発明による単語音声認識装置の一実施例を示す構成図である。図中、斜線で示す部品以外は従来装置の構成および動作と同一であるのでその説明を省略する。

認識モード切替スイッチ16の点線位置を介して順次登録音節数メモリ18に記憶される。一連の単語の登録動作が終了すると、スイッチ16は実線位置に切換え認識動作を開始させる。なお、スイッチ16, 7は連動して切換え動作を行う。従来の方法と同じように、入力単語音声の圧縮パターン6はスイッチ7の点線位置を介して登録パターンメモリ9に格納される。したがって、認識モードにおいて、入力単語音声の圧縮パターン6は入力パターンメモリ8に一時記憶される。また、その音節数15は入力音節数メモリ17に一時記憶される。2個のメモリ8, 17は音声入力毎に更新され書き換えられる。

入力音節数19および登録音節数20は、音節数比較器21に導入される。比較器は両者を比較し、一致したときだけ判定制御信号22を発生し、認識判定回路12の動作を指示する。即ち、長時間を要するパターン照合動作は、入力単語音声と同一の音節数を有する登録単語のみに限定される。したがって、音節数の異なる登録単語パターンと

の照合動作はスキップされ、認識処理の高速化が達成される。

例えば、1, 2, 3, 4音節単語がそれぞれ50語登録されることによつて、合計200語が登録されている単語音声認識装置を仮定する。ここで1音節単語が入力された場合、従来装置では200回のパターン照合動作が必要であるのに対し、この発明による装置では50回の照合動作のみでよく認識処理時間を1/4に短縮することができる。これは2, 3, 4音節単語音声入力時でも同様である。

なお、音節数メモリの容量として1語当たり最高4ビットあればよく、特徴パターンメモリの容量に比べ無視しうる程度である。説明の便宜上、第2図において2分割で示されたモード切換スイッチ7と16、入力パターンメモリ8と入力音節数メモリ17、登録パターンメモリ9と登録音節数メモリ18、ならびに認識判定回路12と音節数比較器21はそれぞれ類似の動作を行うので、実際には一体化した構成も可能である。

認識装置は従来装置のようにパターン照合動作をすべての登録単語に対して行う必要がなく、認識処理の高速化を實現できる。この効果は登録単語数が多い程顕著である。

したがつて、大語彙を取扱う認識装置の処理部に高価な高速プロセッサを使わずとも、安価なマイクロプロセッサ等を利用することができる。さらに、この発明による単語音声認識装置は、従来装置に音節計数回路、比較器および若干のメモリを追加すればよく、極めて容易かつ経済的に實現することも可能である。

4、図面の簡単な説明

第1図は従来の単語音声認識装置の一例を示す構成図である。第2図はこの発明による単語音声認識装置の一実施例を示す構成図である。

図において、1はマイクロホン、2は音声波形、3は音声分析・特徴抽出回路、4は特徴パターン、5は特徴パターン圧縮回路、6は圧縮パターン、7, 16は学習/認識モード切換スイッチ、8は入力パターンメモリ、9は登録パターンメモリ、

ところで、もし音節数の計数が誤つた場合、誤認識を生ずる可能性がある。計数された音節数と同一の登録パターン群との照合結果が思わしくないとき(すなわち類似度が所定のしきい値以下のとき)照合範囲を拡大し、前記音節数前後の該当登録パターン群と照合比較し、最大の類似度をもつ登録単語を選択することも可能である。

別の方法として、当初から計数された音節数およびその前後に相当する登録パターン群との照合動作を行う方法がある。この方法によれば、音節数の計数誤りによる誤認識を補正することが可能である。

さらに、この発明による処理時間の短縮効果は、特徴パターン圧縮を行わずダイナミックプログラミング手法により非等長パターン間の照合動作を行う装置にも適用可能である。また、対象話者は特定人または不特定人を問わない。その他音声以外の音響信号や画像信号等の認識処理の高速化にも適用しうる。

以上説明したごとく、この発明による単語音声

10は入力パターン、11は登録パターン、12は認識判定回路、13は認識結果、14は音節計数回路、15は計数音節数、17は入力音節数メモリ、18は登録音節数メモリ、19は入力音節数、20は登録音節数、21は音節数比較器、22は判定制御信号を示す。

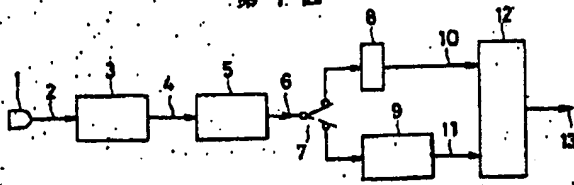
代理人 萬 野 信 一 (外1名)

BEST AVAILABLE COPY

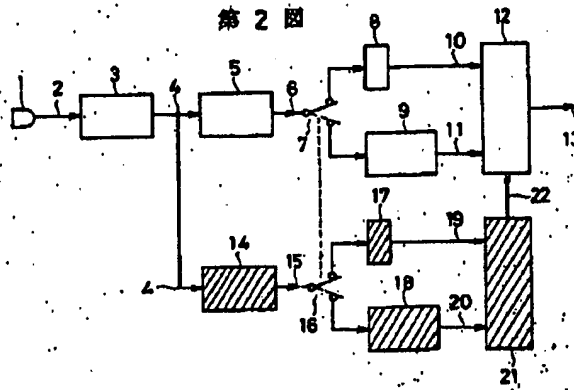
昭和56年7月2日



第1図



第2図



特許庁長官殿

1. 事件の表示 特願昭 55-163148号

2. 発明の名称

音声認識装置

3. 補正をする者

事件との関係

特許出願人

住所
名称 (601)

東京都千代田区丸の内二丁目2番3号

三菱電機株式会社

代表者 池田 謙 三

片山 仁八郎

4. 代理人住所

東京都千代田区丸の内二丁目2番3号

三菱電機株式会社内

弁理士 野 信 一

(特許第2031273/3421号)

56.7.1

(1)

正

BEST AVAILABLE COPY

5. 補正の対象

明細書の特許請求の範囲の欄および発明の詳細

な説明の欄

6. 補正の内容

(1) 特許請求の範囲を別紙のとおり。

(2) 明細書第4頁第8行の「音声の音声」を「音声の音節」に訂正する。

以上

2. 特許請求の範囲

(1) 入力単語音声と登録単語音声を比較し入力単語音声を識別する認識手段、

前記入力単語音声の音節数を検出する検出手段、

前記登録単語音声の音節数を記憶する記憶手段、および

前記入力単語音声の音節数と上記登録単語音声を比較し、認識動作において入力単語音声と同一または近似の音節数をもつ登録単語のみを照合対象とする比較手段を備えた音声認識装置。

(2) 前記認識手段は、音声分析・特徴抽出回路と特徴パターン圧縮回路を含む特許請求の範囲第(1)項記載の音声認識装置。

(3) さらに登録モードにおいて前記検出手段を動作させ認識モードにおいて前記比較手段の動作の結果入力単語音声と登録単語音声が同一または近似の音節数を有するとき前記認識手段を動作させる登録/認識モード切替手段を備えた特許請求の範囲第(1)項記載の音声認識装置。